Journal of Public Economics 199 (2021) 104441

Contents lists available at ScienceDirect

Journal of Public Economics

journal homepage: www.elsevier.com/locate/jpube

The short- and long-run impacts of secondary school absences ☆

Jing Liu [a,*], Monica Lee [b], Seth Gershenson [c]

[a] College of Education, University of Maryland College Park, United States
[b] Annenberg Institute, Brown University, United States
[c] School of Public Affairs, American University and IZA, United States

ARTICLE INFO

Article history:
Received 11 March 2020
Revised 26 April 2021
Accepted 3 May 2021

Keywords:
Student absences
Achievement gaps
Education production function

ABSTRACT

We provide novel evidence on the causal impacts of student absences in middle and high school on state test scores, course grades, and educational attainment using a rich administrative dataset that tracks the date and class period of each absence. We use two similar but distinct identification strategies that address potential endogeneity due to time-varying student-level shocks by exploiting within-student, between-subject variation in class-specific absences. We also leverage information on the timing of absences to show that absences that occur after the annual window for state standardized testing do not affect test scores, providing a further check of our identification strategy. Both approaches yield similar results. We find that absences in middle and high school harm contemporaneous student achievement and longer-term educational attainment: On average, missing 10 classes reduces math or English Language Arts test scores by 3–4% of a standard deviation and course grades by 17–18% of a standard deviation. 10 total absences across all subjects in 9th grade reduce both the probability of on-time graduation and ever enrolling in college by 2%. Learning loss due to school absences can have profound economic and social consequences.

1. Introduction

There is an emerging consensus that student attendance is both a critical input and an intermediate outcome of the education production function. The U.S. Department of Education recently called chronic absenteeism, defined as missing at least 10% of school days, "a hidden educational crisis."[1] Accordingly, education policy-makers are increasingly incorporating student attendance into accountability measures used to gauge schools' and teachers' performance, most notably via the *Every Student Succeeds Act* (ESSA), which has renewed interest in interventions to reduce student absenteeism (Bauer et al., 2018; Gottfried and Hutt, 2019). At the same time, growing evidence shows that a variety of school-based inputs and interventions including effective teachers, small classes, and information-based nudges, can reduce absenteeism (Bergman and Chan, 2021; Gershenson, 2016; Liu and Loeb, 2021; Rogers and Feller, 2018; Tran and Gershenson, 2021).

Heightened policy and research interest in student absenteeism is prefaced on the well-documented correlation between students' absences and educational outcomes representing a causal relationship. While it is intuitive to assume a causal relationship, identification remains a persistent challenge (Jacob and Lovett, 2017). The main as yet unresolved threat to identification is the likelihood that unobserved, time-varying, student-level shocks, such as illness or a family emergency, confound existing estimates by affecting students' attendance and their academic performance. For example, several studies use various combinations of student, family, school, teacher, and classroom fixed effects (FE) to control for the endogeneity of absences. While these studies represent an improvement over prior work, none specifically controls for time-varying, student-specific shocks (Aucejo and Romano, 2016; Gershenson et al., 2017; Gottfried, 2009; Gottfried, 2011). Another set of papers seek to identify causal estimates by using variation

☆ The authors thank Mingyu Chen, Thomas Dee, Matthew Kraft, Susanna Loeb, Nicholas Papageorge, Yang Song, Xiaoyang Ye, Di Xu, and participants in the 2019 Association for Education Finance and Policy (AEFP) conference, 2019 Association for Public Policy and Management (APPAM) fall conference, and 2020 American Economic Association (AEA) annual meeting for helpful discussions. Jessica Horst provided excellent research assistance. The authors acknowledge financial support from The Spencer Foundation, and from the Institute of Education Sciences (R305B090016) to the Board of Trustees of the Leland Stanford Junior University. The data were provided via in-kind support from a local educational agency using a data sharing agreement that requires agency review of the findings of the research prior to publication. Opinions reflect those of the authors and not necessarily those of the funding agencies or the data-sharing partner.

* Corresponding author.
E-mail addresses: jliu28@umd.edu (J. Liu), monica_lee@brown.edu (M. Lee), gershens@american.edu (S. Gershenson).

[1] U.S. Department of Education. Chronic absenteeism in the nation's schools. From: https://www2.ed.gov/datastory/chronicabsenteeism.html

https://doi.org/10.1016/j.jpubeco.2021.104441
0047-2727/

induced by variables such as distance to school, snowfall, and flu cases, to instrument for student absences (Aucejo and Romano, 2016; Goodman, 2014; Gottfried, 2010); these studies find significant effects of absences on achievement, though whether the relevant exclusion restrictions are valid is debatable and, in any case, the resulting local average treatment effect estimates are not necessarily the parameters of broad policy interest, in the sense that absences induced by these instruments might be immutable.

The dearth of credible evidence on how student absences affect academic performance, particularly in middle and high school, is troubling because how, when, and where student absences affect academic performance has important implications for the design and targeting of interventions, the consequences of absence-based accountability policies, and the role of student absences in contributing to socio-demographic gaps in educational outcomes.[2] Moreover, middle and high school students are at a critical developmental stage as they prepare for the transition into college and young adulthood, so understanding the causes and consequences of absences during this time is essential for efforts to reduce high school dropout rates and increase college readiness and enrollment.

The current study contributes to the literature on the impacts of student class-absences by overcoming the threat posed by unobserved student-year shocks using two similar identification strategies that exploit between-subject differences in students' annual absences. We do so by using a decade's worth of administrative data from a large and diverse urban school district in California that include the date and class period of all middle and high school student absences. Our preferred model extends typical value-added models of the education production function by using total absences (across subjects) to proxy for year-specific student shocks and then estimates the effect of subject-*j* absences on achievement in subject *j*. A related approach that stacks the data by subject and conditions on student-by-year FE yields similar results: ten absences reduce math and ELA achievement by 3 to 4% of a test-score standard deviation (SD), which are both practically and statistically significant. However, they are only 50 to 60% as large as estimates that do not account for possibly endogenous student-year shocks, which highlights the utility of our approach and suggests that existing estimates might similarly overstate the effects of absences.

Finally, we provide relatively novel evidence on the causal effects of high school student absences on longer-term educational outcomes. Specifically, ten absences in the 9th grade reduce both the probability of high school graduation and of ever enrolling in college by 2%. These results cross-validate the main test score results and show that absences not only harm contemporaneous performance on state tests and in specific courses but have long lasting consequences for educational attainment, which is what ultimately matters. These estimates are arguably causal, as we apply selection-on-observables bounding methods that show that an implausibly large degree of sorting on unobservables is needed to explain away the estimated effects (Altonji et al., 2005; Oster, 2019).

The paper proceeds as follows. Section 2 describes the administrative dataset used in our analyses. Section 3 describes our identification strategy. Section 4 presents the main results on absences' impacts on academic achievement and associated tests for heterogeneous effects. Section 5 presents a number of sensitivity analyses and suggestive tests of the main identifying assumptions, as well as estimated effects on course grades and a placebo exercise that distinguishes between absences that occurred before, during, and after the standardized testing window. Section 6 examines the long-run effects of ninth- and tenth-grade absences on high school graduation and college enrollment. Finally, Section 7 concludes with a summary and discussion of policy implications.

2. Data

2.1. Administrative Data

Our main analyses use administrative data from a large urban school district in California from the school years 2002–2003 through 2012–2013.[3] It contains information on student attendance, student and teacher demographics, and students' academic performance, including scores on state-mandated tests and course-level grades.

The attendance dataset is unique in that it contains students' attendance records for each course on each day, along with whether or not an absence was formally excused. During the time-frame of the current study, teachers used paper scantron sheets to mark students as absent, tardy, or present in each class period. Absences were marked as excused if the school received a phone call from a parent or guardian providing reasons for the absence; otherwise, the absence was marked as unexcused. Prior studies using the same data to examine attendance gaps by socioeconomic status (Whitney and Liu, 2017) and teachers' impacts on student attendance (Liu and Loeb, 2021) validate the attendance data's accuracy.

These data are ideal for the current study for several reasons. First, rich class-level attendance data that include the class period, day, and course are rarely available to researchers. Nearly all existing attendance studies use full-day absences to measure total absences. Since part-day absences account for more than half of total class-absences in secondary school and are mostly unexcused (Whitney and Liu, 2017), disregarding part-day absences may result in considerable measurement error, which may bias estimates of the impact of absences on student achievement, especially when part-day absences are nonrandomly distributed across the student population. In addition, such nuanced data not only allow us to compute the total class absences a student has for a specific class but also provide flexibility to code absences based on the exact date they occurred, a key feature that allows us to verify our identification strategy. Lastly, the district has a large and diverse student body, which provides the power and variation necessary to conduct subgroup analyses.

2.2. Constructing the analytic sample

We combine three databases to construct the analytic sample. First, we match the attendance data to student course-taking data and identify the corresponding class, subject area, and end-of-course grades. The specificity of the attendance data allows us to distinguish between full-day absences when a student misses every single class and part-day, or course-specific absences. The latter facilitates the identification of subject-specific absences.

[2] The majority of credible evidence on the impact of absences comes from the elementary school setting; a notable exception is Kirksey (2019), who uses student fixed-effects models to show that high school students' full-day absences harm achievement in a small urban district in California. Interventions designed to reduce student absences are similarly disproportionately focused on the early grades (Bauer et al., 2018). Notable exceptions here are Rogers and Feller (2018), who found that a light-touch information intervention reduced chronic absenteeism rates by about 10% in both primary and secondary grades, and Bergman and Chan (2021), who found that text messages to the parents of middle and high school students about their grades and attendance increased student attendance by 12%.

[3] While the external validity of any analysis of a single school district is a potential concern, the rich course-level and long-run data available here, as well as the large student population, make these data ideal for the current study. Moreover, despite the somewhat unique demographic composition of the student body, this large urban district faces many of the same challenges faced by other large districts.

We focus our analysis on math and English language arts (ELA), as these two subjects are consistently tested across all grade levels in state-mandated exams.[4] We exclude remedial and tutoring courses, study halls, and courses for English Language Learners (ELL) because the instructional content of these courses is not covered in state standardized tests.[5] This allows us to observe the total number of absences, end-of-course grade, and state exam scores for each student in each school year and subject. Second, we merge in a rich set of demographic variables, including race/ethnicity, gender, ELL status, special education status, disability status, and residential census tracts (which provide a measure of socioeconomic status).

In addition, we categorize absences as ones that occurred before, during, or after the annual state standardized testing window. We refer to the testing window and not the test date for two reasons. First, before 2009, we do not know the exact dates that schools first administered the test; we only know that California required all standardized tests to be administered within a 21-day window centered on the day on which 85% of yearly instruction is complete.[6] Second, even when we do observe the exact date that a school began testing, we do not observe the dates that individual students took the test, as not every student in a school took the test on the same day for many reasons. Thus, while we can clearly identify absences that occurred prior to the testing window–and thus before the test was taken–there is some ambiguity as to whether absences late in the year occurred before or after the student took the test. Accordingly, in our baseline model, we only use absences that took place prior to the state-sanctioned testing window to estimate the effect of absences. We incorporate testing window information as a falsification check in Section 5, given that absences after the test date should not affect test scores.

Lastly, we augment the dataset to include several long-run outcomes. We observe whether a student graduated from the district "on time" (i.e., four years after initial enrollment in high school).[7] We also observe students' post-secondary enrollment data, which the district obtained from the National Student Clearinghouse, a nonprofit organization that provides degree and enrollment verification for more than 3,300 colleges and 97% of students nationwide (Dundar and Shapiro, 2016). Such data is available through the end of 2016, covering slightly more than 55% of students in our sample.[8] We incorporate these long-term outcomes, including on-time high school graduation, immediate college enrollment, whether a student ever enrolled in college, and whether it is a four-year or two-year college, to evaluate whether absences in secondary school have effects above and beyond their immediate impact on student achievement.

2.3. Descriptive statistics

The main analytic sample consists of over 70,000 students, whose average characteristics are summarized in Table 1. The district is racially diverse: About 47% of students are Asian; nearly one-quarter are Hispanic; and about 13% are Black. Given this diversity, it is not surprising that a quarter of the students observed are flagged as ELLs. Average achievement hovers slightly above zero across both subjects and score type.[9] Lastly, over half of all students attend some type of postsecondary institution after high school graduation.

Table 1
Sample Means

	Mean	SD
A. Student Characteristics		
Female	0.48	
White	0.09	
Black	0.13	
Hispanic	0.22	
Asian	0.47	
Other Race	0.09	
Special Education	0.11	
Gifted	0.25	
ELL	0.25	
Disabled	0.05	
Chronically Absent	0.08	
B. Short-Term Outcomes (Student-Year Level)		
Math Course Grade	0.02	(0.96)
ELA Course Grade	0.00	(0.97)
Math Test Score	0.05	(0.99)
ELA Test Score	0.04	(0.98)
C. Long-Term Outcomes		
Graduated High School	0.65	
Enrolled in College Immediately after High School	0.50	
Ever Enrolled in College	0.57	
Ever Enrolled in 4 Year	0.41	
Ever Enrolled in 2 Year	0.40	
Student-Year Observations	210,380	
Student-Year-Subject Observations	383,534	
Number of Unique Students	72,161	
Number of Courses	11,906	
Number of Schools	43	

Note: Data include 7th- to 11th-graders from school years 2002–2003 to 2012–2013. Panel B summarizes characteristics at the student-year level, and Panels A and C summarize characteristics at the student level. Students are flagged in a given category for Special Education, Gifted, ELL, Disabled, or Chronically Absent if they are ever flagged for that category across any of the years enrolled. Course grades are standardized at the school-year-subject-course level and test scores are standardized at the school-year-subject level.

In Fig. 1 we observe that the distributions of total absences in math and ELA classes are nearly identical. In both subjects, approximately one-fifth of students have zero absences while a long tail indicates a nontrivial number of students who accumulate 18 or more absences, a common definition of chronic absenteeism across most states.

Table 2 reports average absences and achievement across both subjects by student subgroup. Absence patterns are similar across subjects. On average, students miss about eight classes prior to the start of the state-mandated testing window. Students miss almost two days of school during the testing window and another 1.5 days after. This dynamism presents an opportunity for a falsification test of our main finding, which we discuss in Section 5.2. Within-student SD of absences are relatively large compared to the overall SD, suggesting that about half of the variation in absences occurs within, as opposed to between, students. Table 2 also highlights racial gaps in attendance and achievement in both subjects. Hispanic and Black students miss two and three times as many classes, respectively, as white and Asian students and this racial gap exists throughout the school year.

[4] Students in grades 2 through 11 were required to take state standardized tests during this time.

[5] We also exclude students who enrolled in multiple math or ELA classes in the same semester, as this subgroup of students makes up less than 5% of the overall sample and are not comparable to the general student population. These students tend to be special education students, students with individualized plans for learning, or students who are otherwise on an alternative pathway to high school completion.

[6] In the last two years (2011–2012 and 2012–2013), the state expanded the testing window to 25 days, which we account for in our categorization of absences.

[7] The district does not track graduation information for students who do not meet graduation requirements, drop out, or move away from the district prior to graduation. Based on the district's suggestion, we code students who should have graduated from high school prior to 2015 but whose graduation status is missing as not having graduated on time.

[8] There is no college enrollment information for students who do not officially graduate or drop out from the district. To be consistent, we code these students as not enrolling in college.

[9] This may be the case due to the exclusion criteria we mention above, where many special education students and those who would otherwise perform poorly are excluded from the analytic sample.

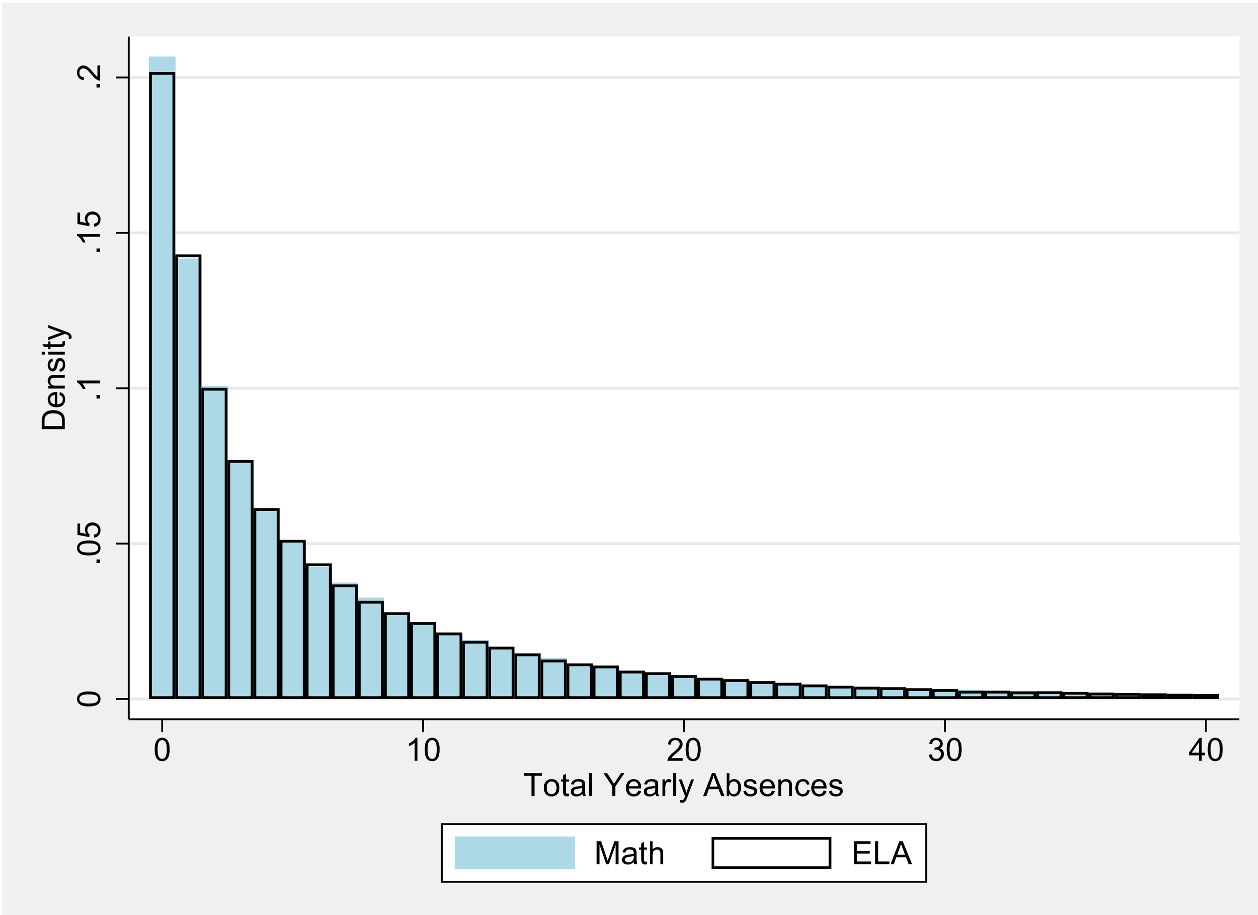

Fig. 1. Distribution of Absences by Subject. Note: Observations are counts of yearly student absences in math or ELA classes between 2002–2003 to 2012–2013. Data are trimmed at 40 absences to show the bulk of variation.

3. Identification Strategy

Our novel approach to isolating the causal effects of secondary school absences on academic achievement builds on a basic value-added model of the education production function, in which student absences are a contemporaneous input (Aucejo and Romano, 2016; Gershenson et al., 2017). Specifically, we model student i's standardized end-of-year test score in subject j in year t as a function of possibly time-varying student characteristics (X_{ijt}) and their total class absences in the subject (a_{ijt}). Following the value-added literature, X_{ijt} includes lagged test scores that control for students' sorting into classrooms and their prior academic history (Chetty et al., 2014). This model is specified as:

$$Y_{ijt} = \beta X_{ijt} + \gamma a_{ijt} + \mu_{ijt}, \tag{1}$$

where X includes lagged test scores, ELL and special education status, demographic background, and census-tract-by-year indicators to control for socio-economic status and neighborhood effects and μ represents the unobserved determinants of student achievement.

However, unlike in the case of teacher effects or other school inputs that are assigned at the start of the school year, OLS estimates of Eq. (1) are likely biased due to the fact that absences (a_{ijt}) are both an intermediate output *and* an input of the education production function, which might be affected by the same idiosyncratic shocks that affect test scores. Specifically, student-specific, time-varying shocks are a threat to the identification of the causal effect of absences that neither lagged test scores nor student fixed effects adequately control for. For example, an illness or family emergency specific to year t could jointly affect student i's absences and achievement. Existing research on the causal effects of student absences acknowledges, but has yet to fully address, this concern.

We address this identification challenge by leveraging between-subject variation in student absences and assuming that such variation is conditionally random. We do so using two related but distinct approaches. The first uses total annual absences in math and ELA to proxy for the year-specific unobserved shock and is described in Section 3.1. The second stacks the data across subjects so there are two observations per student-year and estimates a model that conditions on student-by-year FE; this approach is described in Section 3.2.

The two approaches are similar in that they rely on the same type of identifying variation, make the same key identification assumption, and yield similar results. However, each approach also makes some additional, unique assumptions. The general identification strategy, as well as the subtle differences between our two approaches, are motivated by specifying the absence production function.

Specifically, we model a_{ijt} as a function of students' time-invariant subject-specific preferences (π_{ij}), subject-year effects (α_{jt}), student-year shocks (θ_{it}), and an idiosyncratic error:

$$a_{ijt} = \pi_{ij} + \alpha_{jt} + \lambda\theta_{it} + e_{ijt}, \tag{2}$$

where $e_{ijt} \perp \theta_{it}$ by definition. This model suggests that to causally identify the effects of absences in an education production function like that specified in Eq. (1), we need to address biases potentially

Table 2
Absences by Subject and Time.

	All Students		White/Asian		Hispanic		Black	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
A. Math								
Absences								
Yearly Total	11.1	(16.7) [8.4]	7.0	(12.8) [7.0]	17.4	(19.5) [10.5]	22.5	(22.1) [10.8]
Before Testing Window	8.0	(12.4)	5.0	(9.5)	12.6	(14.6)	16.1	(16.5)
During Test Window	1.8	(3.3)	1.1	(2.5)	2.9	(3.9)	3.8	(4.4)
After Test Window	1.5	(2.5)	1.0	(2.0)	2.2	(2.9)	3.0	(3.4)
Achievement								
Test Score	0.04	(0.96)	0.36	(0.90)	−0.53	(0.74)	−0.68	(0.75)
End of Course Grade	−0.00	(0.96)	0.15	(0.95)	−0.22	(0.95)	−0.39	(0.91)
Student-Year Observations	190,972		109,724		39,803		20,100	
Unique Students	70,525		38,995		15,116		8,349	
B. ELA								
Absences								
Yearly Total	11.1	(16.8) [8.4]	6.7	(12.2) [6.5]	17.4	(19.7) [10.8]	23.0	(22.8) [11.9]
Before Testing Window	8.0	(12.5)	4.8	(9.1)	12.6	(14.7)	16.6	(17.0)
During Test Window	1.8	(3.3)	1.1	(2.4)	2.8	(3.9)	3.8	(4.4)
After Test Window	1.5	(2.5)	0.9	(1.9)	2.2	(2.9)	3.0	(3.5)
Achievement								
Test Score	0.13	(0.95)	0.42	(0.84)	−0.35	(0.87)	−0.63	(0.88)
End of Course Grade	0.00	(0.97)	0.17	(0.92)	−0.26	(0.99)	−0.41	(0.95)
Student-Year Observations	192,562		110,194		37,538		22,540	
Unique Students	67,432		36,810		13,975		8,724	

Note: SD in brackets indicate within-student SD. Student-year observations and number of unique students are specific to each subject (Math in Panel A, ELA in Panel B). Statistics of students in Other Race category are omitted from display. Data include 7th- to 11th-graders from the school years 2002–2003 to 2012–2013. The testing window is an estimate of the test window based on the California Education Code and does not necessarily reflect the exact test dates, which might include make-up periods at the discretion of each school. .

induced by π_{ij}, α_{jt}, and θ_{it}. It is straightforward to account for π_{ij} with either lagged test scores or student FE. We can also easily account for α_{jt} by conditioning on classroom FE (Gershenson et al., 2017).[10]

However, the student-year shock θ_{it} is more problematic, as the unit of analysis in Eq. (1) is the student-year. Below, we present two distinct but related approaches to controlling for θ_{it} that exploit the fact that student absences vary across subjects. The subtle difference between the two approaches is in how λ maps the idiosyncratic shock into absences. The lack of subscripts on λ in Eq. (2) indicates that this mapping is the same in both subjects, though we later relax this assumption. Intuitively, a constant λ means, for example, that a stressful family emergency in year t would not cause a larger increase in absences in math than in ELA, even if the student had a pre-existing preference for ELA classes, as given by π_{ij}. Alternatively, if λ varies by subject, the student might respond to the same family emergency by increasing absences differently in different subjects. Whether we assume a constant or subject-specific λ leads to two distinct identification strategies.

Importantly, both approaches require that students have no time-varying, subject-specific preferences that influence both absences and achievement. For example, this assumption would fail if a student had better performance and fewer absences in class-j because her friends were in the class or she felt a special bond with the teacher. These sorts of time-varying, student-specific subject preferences would enter the idiosyncratic error e_{ijt} in Eq. (2) and are distinct from the issue of how λ maps shocks into outcomes. The plausibility and some checks of both assumptions are discussed below and in Section 5.[11]

3.1. Proxy approach

If λ in Eq. (2) is the same for all subjects and there is no year-specific preference for one subject over the other, the difference between a_{ijt} and $a_{i,-j,t}$ is conditionally random.[12] This suggests using total annual absences across both subjects (A_{it}) as a proxy for θ_{it}. Adding Eq. (2) across subjects yields

$$\begin{aligned} A_{it} &= a_{ijt} + a_{i,-j,t} \\ &= (\pi_{ij} + \pi_{i,-j}) + (\alpha_{jt} + \alpha_{-j,t}) + 2\lambda\theta_{it} + (e_{ijt} + e_{i,-j,t}), \end{aligned} \tag{3}$$

where it is straightforward to solve Eq. (3) for θ_{it} and recognize that A_{it} is a potential proxy for the unobserved student-year shock in a typical education production function, as the π and α terms would be controlled for by the lag scores and classroom FE, respectively, which are common to such models. By inserting Eq. (3) into Eq. (1), formally, we estimate

$$Y_{ijt} = \beta X_{ijt} + \gamma a_{ijt} + \delta A_{it} + u_{ijt}, \tag{4}$$

[10] Conditioning on classroom FE has additional benefits in the secondary school setting because they account for otherwise unobserved differences between classrooms such as regular versus block schedules (Rice et al., 2002), "tracking"-based differences in rigor (Jackson, 2014), and the physical location and time that classes meet (Carrell et al., 2011; Heissel and Norris, 2018; Pope, 2016).

[11] In a descriptive analysis of the same administrative data used here, Whitney and Liu, 2017 compare the predictors of partial- and full-day absences and find little, if any, evidence of time-varying, subject-specific preferences in class skipping. Specifically, only about half of full-day absences are unexcused, while more than 90% of part-day absences are unexcused. The main source of variation in part-day (subject-specific) absences is the class meeting time and not subject, which is controlled for by classroom fixed effects. This provides additional supporting evidence for the main identifying assumption in the current study: that remaining between-subject variation in students' annual absences is as good as random.

[12] This is similar to the identifying assumption in twins-based research designs (Ashenfelter and Krueger, 1994; Bound and Solon, 1999).

where u is a composite error term that includes e_{ijt} and $e_{i,-j,t}$.[13] We estimate Eq. (4) by OLS separately for math and ELA achievement, where A is the same in both models.[14] This identification strategy effectively compares two students who have the same number of total absences but one has an additional math absence (thus one fewer ELA absence) than the other, holding other variables constant.[15] A causal interpretation of these estimates makes two assumptions on top of the "constant λ" assumption discussed above and the usual value-added model assumptions regarding the effectiveness of controlling for lagged test scores and classroom FE.

First, conditional on a_{ijt}, X, and θ_{it}, it must be that $a_{i,-j,t}$, and therefore A_{it}, does not directly affect Y_{ijt} and is thus available as a proxy. This is the textbook "redundancy assumption" of proxy plug-in solutions to omitted variables bias (Wooldridge, 2010). This rules out spillover effects of absences in one subject on achievement in the other; while this is unlikely to hold exactly, we argue that such spillovers are likely negligible and provide supporting empirical evidence of this in Section 5.

Second, the ratio of a_{ijt} to $a_{i,-j,t}$ must be conditionally random. Technically, this is because the idiosyncratic errors of Eq. (3) enter the composite error of Eq. (4) and is analogous to the textbook assumption of a valid proxy. Here, it intuitively means that students do not have time-varying subject preferences that directly affect absences and achievement. Relative to students' time-invariant preferences and classroom effects that are common to all students, we argue that time-varying preferences of this sort are likely negligible and provide supporting empirical evidence that this is so in Section 5.

3.2. Stacked model

Alternatively, suppose that the λ in Eq. (2) varies by subject (i.e., the underlying preferences π_{ij} moderate the effect of θ_{it}, or there is an interaction effect between π_{ij} and θ_{it}). For example, a student with an underlying math preference might respond to a stressful family situation in year t by only (or disproportionately) increasing their ELA absences. In this case, A_{it} is no longer a useful proxy because the a_{ijt} and $a_{i,-j,t}$ in Eq. (3) are weighted differently, and since these weights are unknown, we cannot distinguish them from the parameter of interest γ.

Instead, we can directly control for potentially endogenous student-year shocks by stacking the data across subjects so that there are two observations per student-year (math and ELA) and estimating a model that conditions on student-by-year FE. Formally, we estimate

$$Y_{ijt} = \gamma a_{ijt} + \alpha_j + \phi_{it} + u_{ijt}, \tag{5}$$

where α and ϕ are subject and student-year FE, respectively.[16] In contrast to the proxy approach, Eq. (5) is agnostic about the nature of λ in Eq. (2), but requires that the student-year shock affects achievement *the same way in both subjects*. As a practical matter, the student-year FE make any student-year or school-year controls redundant, though the model can accommodate classroom FE to account for any unobservables (e.g., teacher quality and peer composition) that are constant at the classroom level.

Importantly, a causal interpretation of the stacked model's estimates requires the same assumption of no time-varying, subject-specific preferences made by the proxy model. Intuitively, this is analogous to the assumption of no time-varying heterogeneity in standard panel data models, and once again requires that the ratio of a_{ijt} to $a_{i,-j,t}$ in a given year is conditionally random. The results presented in Section 4.1 show that estimates of γ in Eqs. 4 and 5 turn out to be qualitatively similar, suggesting that the subtly different identifying assumptions regarding λ are unimportant. A thorough assessment of the plausibility of all the identifying assumptions is presented in Section 5.

4. Student absences and academic achievement

4.1. Main results

Table 3 reports estimates of the education production functions specified in Eqs. (4) and (5). Columns (1) through (3) report versions of the "proxy" model (Eq. 4) for math achievement. Column (1) presents estimates from a relatively standard, fully-specified value-added model that conditions on linear and quadratic lagged test scores in both math and ELA, observed student characteristics, lagged GPA, and classroom and neighborhood-by-year FE. However, this model fails to adjust for unobserved, student-year shocks. Thus, the relatively large point estimate of -0.08, which suggests that ten math absences reduce achievement by 8% of a test-score SD, is likely biased upwards by the presence of student-year unobserved shocks that jointly determine absences and achievement.

Column (2) reports estimates from our preferred model, which proxies for unobserved student-year shocks with A_{it}, a student's total annual absences across both subjects. Doing so cuts the point estimate on math in half, suggesting that the causal effect of ten math absences is about 4% of a test-score SD. This significant decrease suggests that the estimate in Column (1), and many existing estimates of the effect of student absences that rely on lagged test scores or student-FE strategies, are biased upward by unaccounted-for idiosyncratic shocks. Column (3) replaces the lagged test scores and time-invariant student characteristics with student FE, which leaves the point estimate virtually unchanged, and reassures us that selection into classrooms and pre-existing, time-invariant predispositions for being absent are adequately controlled for.

Columns (4)-(6) estimate the same three model specifications for ELA. The results here follow a similar pattern: in all three models there is a significant and negative relation between ELA absences and ELA test scores, though conditioning on A_{it} reduces the point estimate by about 33%. The preferred estimate in column 5 shows that ten ELA absences reduce ELA scores by about 4% of a test-score SD, which is quite similar to the estimated effect of math absences on math achievement. Interestingly, the naive estimates that do not account for idiosyncratic shocks in Columns (1) and (4) show a larger effect of absences on math than ELA scores, which is consistent with past research (Gershenson et al., 2017); however, this difference vanishes upon controlling for idiosyncratic shocks, which once again highlights the importance of accounting for time-varying, student-specific shocks that jointly affect absences and achievement.

Finally, Columns (7) and (8) of Table 3 estimate the stacked student-year FE model specified in Eq. (5). The estimates in Column (8) also condition on classroom FE, though this proves inconsequential: both models indicate that ten absences reduce achievement by about 3% of a test-score SD. These estimates are

[13] The classroom FE in X subsume the year, school, and teacher FE, as well as observed classroom characteristics such as class size and socio-demographic composition, that are typically included in value-added models.

[14] In practice, estimated γ are the same if we control for the other subject's absences $a_{i,-j,t}$ instead of A_{it}.

[15] Standard errors are clustered at the school level, which yields more conservative standard errors than clustering at the classroom or student level.

[16] To be true to Eq. (2), the stacked model could replace α_j with student-by-subject FE to explicitly control for students' time-invariant preferences for subjects (π_{ij}). However, doing so would require restricting the analytic sample to students who took both math and ELA classes in the same year for at least two years, which is a nontrivial restriction in the high school sample. We discuss results from both specifications in Section (4.1).

Table 3
Main results: the impact of absences on test scores.

	Subject Specific Model						Stacked Model	
	Math			ELA			Math + ELA	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Subject Absences	−0.082**	−0.042**	−0.041**	−0.064**	−0.040**	−0.044**	−0.026**	−0.029**
	(0.006)	(0.010)	(0.007)	(0.005)	(0.007)	(0.005)	(0.004)	(0.004)
Math + ELA Absences		−0.025**	−0.021**		−0.015**	−0.012**		
		(0.006)	(0.004)		(0.004)	(0.002)		
Classroom FE	X	X	X	X	X	X		X
Student Controls	X	X		X	X			
Neighborhood-year FE	X	X		X	X			
Student FE			X			X		
Student-Year FE							X	X
R^2	0.720	0.721	0.863	0.754	0.754	0.875	0.842	0.841
Observations	112,711	112,711	176,233	117,445	117,445	181,010	333,624	333,624

Notes: Each column indicates a separate model, with the outcome being math or ELA test scores. Columns (1) to (3) present results from regressions estimating impact of math absences on math test scores, while Columns (4) to (6) present results from regressions estimating impact of ELA absences on ELA test scores. Columns (7) and (8) present results that are from data that stack math and ELA observations into a student-year-subject dataset. All coefficients are multiplied by 10 to ease interpretation. Student-level controls are included in Columns (1), (2), (4), and (5) and they consist of both linear and quadratic lagged math and ELA test scores (standardized), lagged total absence rate, lagged total suspension days, race, gender, ELL status, disability status, and special education status. Student FE supplant student-level controls in Columns (3) and (6), respectively. As we no longer need to use student test scores in the 6th grade as lagged achievement for 7th graders, the sample sizes in Columns (3) and (6) are bigger than Columns (1), (2), (4), and (5). Classroom- and school-level controls consist of the same set of control variables as the student level. Standard errors are clustered at the school level for columns (1) to (6) and at the student level for columns (7) and (8); they are shown in parentheses. + $p < 0.10$ * $p < 0.05$ ** $p < 0.01$.

slightly smaller, but similar in magnitude, to the preferred subject-specific "proxy" estimates reported in Columns (2) and (5). The similar estimates across the two models suggests that subtle differences in identifying assumptions are empirically unimportant once the main threats to identification have been addressed.[17]

4.2. Heterogeneity

Table 4 augments the preferred "proxy" model specified in Eq. (4) to include interaction terms that allow the effects of absences to vary by certain student and school characteristics. Panel A reports estimates for math and panel B does so for ELA. Column (1) interacts absences with a high-school indicator, as the reasons for absences in middle school might be different than in high school. Moreover, high school absences might involve more missed work that needs to be made up, or that is more difficult to make up at home. In Panel A, the effect of absences is slightly less harmful in high school, though this difference is not statistically significant at traditional confidence levels. Panel B, however, shows that ELA absences are only harmful in high school. Again, this may be because parents are more able to help their children make up missed work in middle school, middle school ELA teachers are more proactive about helping students catch up, or because there is less work to make up following a middle school ELA absence.

Columns (2) and (3) test for demographic differences in the harm of absences. This is motivated by evidence of within-school differences by gender in how students respond to school inputs (Autor et al., 2016) and by race in how students are viewed by teachers (Gershenson et al., 2016). In the case of absences, however, we see no evidence of heterogeneous impacts. Similarly, Column (4) of Table 4 tests for differential impacts by school poverty rate. This is motivated by evidence that economically-disadvantaged elementary school students are absent more often, and marginally more harmed by absences, than their more advantaged peers (Gershenson et al., 2017). However, these interaction terms are imprecisely estimated, and provide no clear evidence of socio-economic differences in the harm of absences.

4.3. Functional form

The models estimated thus far assume that there are neither diminishing nor increasing costs to student absences and that there is no discontinuity at the threshold of being considered chronically absent, which in this district is defined as missing at least 10% of total school days (accruing about 18 absences). We make this simplifying assumption in the baseline model because Gershenson et al. (2017) and Kirksey (2019) found the marginal effects of student absences to be approximately constant. Here, we include quadratic and non-parametric functions of absences in the baseline "proxy" model (Eq. 4) to test whether the same is true in the middle and high school settings of interest in the current study.[18]

Fig. 2 plots fitted regression lines from the linear (baseline), quadratic, and non-parametric models. Panel A does so for math and panel B does so for ELA. The different specifications mostly overlap with each other, suggesting that the marginal effect of absences is approximately linear in both subjects, particularly in ELA. There are slightly increasing costs at higher levels of math absences, though these departures from linearity are not statistically significant. Nor is there any discrete change at the threshold of chronic absenteeism (though chronically absent students perform significantly worse than students with just one or two absences). However, the non-parametric estimates are fairly noisy for higher absence counts, as there are relatively few students in those bins, and suggest slightly larger effects of single-digit absences in math than the linear specification, which provides more conservative estimates of the harmful effects of student absences.

[17] Adding student-by-subject FE to the stacked model estimated in columns 7 and 8 attenuates the estimated effects of absences by about 25%, though they remain strongly statistically significant. Re-estimating the preferred model (Columns (2) and (4)) using the same restricted sample (students with both math and ELA scores in multiple school years) attenuates those estimates by about 14%. The similar reductions in magnitude suggest that this is due to the sample restriction and not a bias due to omitting π_{ij} from the stacked model. Accordingly, we prefer using the unrestricted, more generalizable sample in our main analysis.

[18] We interact the linear and quadratic absence counts with an indicator for having at least one absence to better align with the non-parametric plot.

Table 4
Heterogeneity by student- and school-level characteristics.

	(1)	(2)	(3)	(4)
A. Math				
Subject Absences	−0.057**	−0.042**	−0.040*	−0.046**
	(0.012)	(0.012)	(0.018)	(0.017)
Subject Absences × ...				
High School	0.018			
	(0.015)			
Female		0.001		
		(0.010)		
Black			−0.003	
			(0.025)	
Hispanic			0.020	
			(0.017)	
Other			−0.030	
			(0.027)	
Second Tertile, School Poverty				0.023
				(0.025)
Top Tertile, School Poverty				−0.021
				(0.022)
Observations	112,711	112,711	112,711	112,711
B. ELA				
Yearly Subject Absences	0.001	−0.037**	−0.043**	−0.036**
	(0.015)	(0.009)	(0.011)	(0.011)
Yearly Subject Absences × ...				
High School	−0.048**			
	(0.017)			
Female		−0.004		
		(0.011)		
Black			0.019	
			(0.015)	
Hispanic			−0.001	
			(0.016)	
Other			−0.006	
			(0.020)	
Second Tertile, School Poverty				−0.018
				(0.015)
Top Tertile, School Poverty				0.022
				(0.014)
Observations	117,445	117,445	117,445	117,445

Note: Each column and panel indicate a separate model. All coefficients are multiplied by 10 to ease interpretation. Each model interacts absences with a student- or school-level characteristic. All models control for classroom and neighborhood-by-year FE, student-level controls used in Table 3, and total absences in both math and ELA. The reference group is middle school students for column (1), males for column (2), white students for column (3), and schools that are in the first tertile in student poverty rate for columns (3) and (4). Standard errors are clustered at the school level and shown in parentheses. + $p < 0.10$ * $p < 0.05$ ** $p < 0.01$.

5. Robustness and model validation

5.1. Course grades

Course grades provide an alternative, contemporaneous measure of learning. It is useful to document the impact of absences on another outcome that is not subject to the critiques of standardized test scores (e.g., measurement error and "teaching to the test"). And although grades are a more subjective measure of academic performance than test scores, they are an important outcome in their own right as high school grades tend to be more predictive than test scores of long-term student success (Easton et al., 2017).[19] Moreover, they determine GPA, class rank, and honors, all of which contribute to students' college plans and admissions. To our knowledge there is little credible evidence on how absences affect course grades, though Gottfried (2010) does use an instrumental variables strategy to show an effect of absences on cumulative GPA in the School District of Philadelphia.

Columns (1) and (2) of Table 5 show estimates of the baseline "proxy" model (Eq. 4) for math and ELA course grades, respectively. Course grades are standardized to have mean of 0 and SD of 1, and again we scale the point estimates to reflect the impact of ten absences. Subject-specific absences significantly reduce grades in both subjects by almost 20% of a SD. It is possible that some teachers explicitly factor in attendance in course grades (in violation of district rules), but the effect is too big to be solely a reflection of a mechanical relationship. This suggests that absences are a true impediment to learning with potential long-lasting consequences, and not just missed preparation for end-of-year tests.

5.2. Falsification test

Access to course grades and the exact timing of absences also facilitate a falsification test of our main finding that absences reduce test scores. The idea is that since state mandated tests occur during the approximately four-week testing window that usually starts in week 15 or 16 of the spring semester, absences after the testing window cannot affect test scores but can affect course grades.[20] If post-test absences "affect" test scores, we would worry that our identification strategy is not adequately controlling for selection into absences. Herrmann and Rockoff (2012) and Gottfried and Kirksey (2017) conduct similar falsification tests in

[19] Grades in the district are explicitly based on academic performance and are not allowed to be based on nonacademic factors such as absences.

[20] As discussed in Section 2, we can only impute testing windows, as the exact timing of testing is unknown and may vary across students within a school. We define the start dates using the state education code. A typical testing window starts in week 15 or 16 in the spring semester and lasts until week 19 or 20.

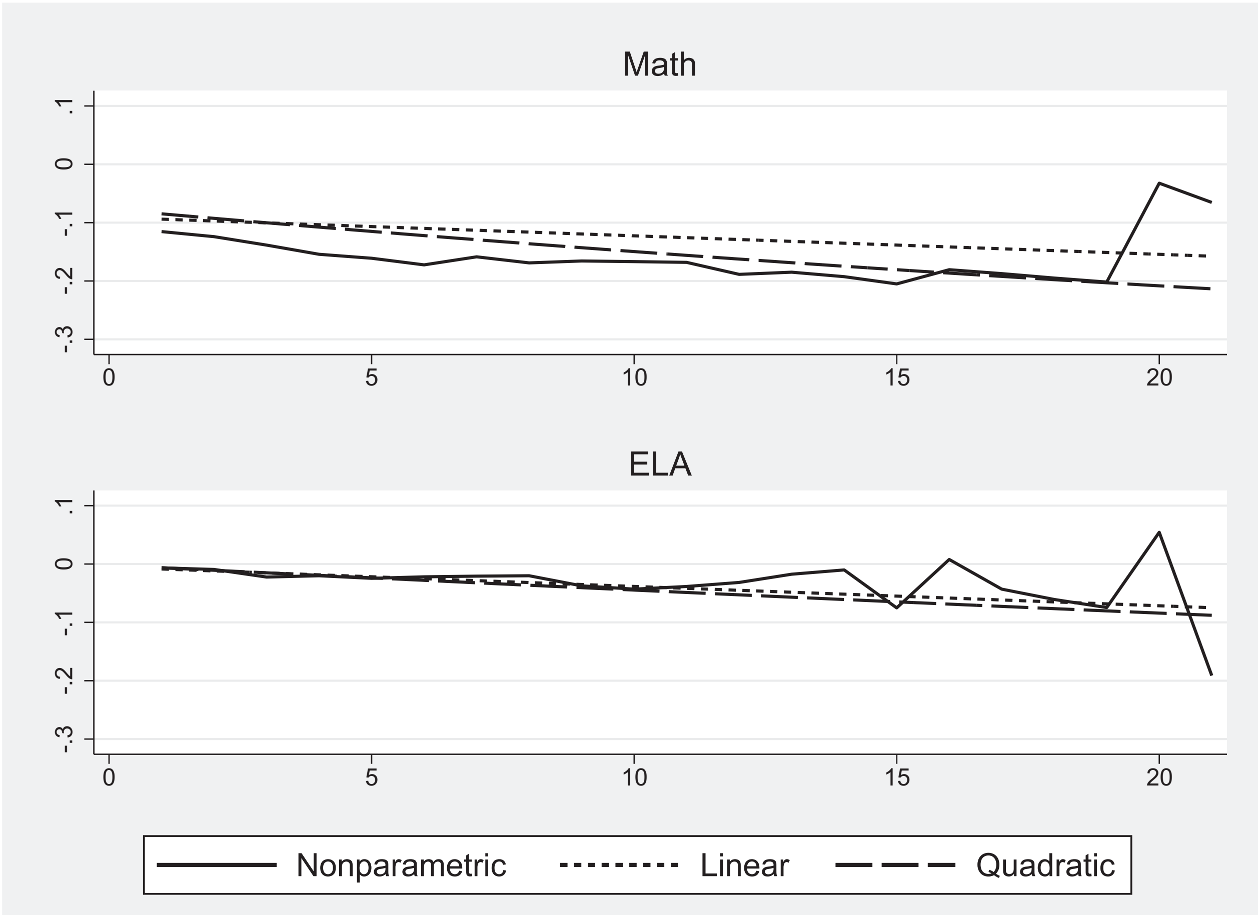

Fig. 2. Linear, Quadratic, and Nonparametric Estimates of the Impact of Absences on Achievement. Note: The graph plots linear, quadratic, and nonparametric models separately. In the nonparametric model, absences beyond 20 are binned as one group.

Table 5
Impact of test window absences on achievement.

	Course Grades				Test Scores	
	Math (1)	ELA (2)	Math (3)	ELA (4)	Math (5)	ELA (6)
Before Test Window						
Subject Absences	−0.169**	−0.180**	−0.086**	−0.110**	−0.017*	−0.017**
	(0.011)	(0.010)	(0.011)	(0.010)	(0.007)	(0.006)
ELA + Math Absences	−0.052**	−0.062**	−0.032**	−0.023**	−0.015**	−0.005+
	(0.007)	(0.008)	(0.008)	(0.006)	(0.004)	(0.003)
During Test Window						
Subject Absences			−0.416**	−0.316**	−0.113**	−0.083**
			(0.056)	(0.035)	(0.020)	(0.026)
ELA + Math Absences			−0.059*	−0.166**	−0.059**	−0.072**
			(0.026)	(0.021)	(0.013)	(0.010)
After Test Window						
Subject Absences			−0.475**	−0.608**	0.031	−0.108
			(0.087)	(0.096)	(0.076)	(0.067)
ELA + Math Absences			−0.197**	−0.197**	−0.112**	−0.019
			(0.047)	(0.051)	(0.037)	(0.035)
Observations	112,325	117,527	112,325	117,527	105,773	106,825
R^2	0.337	0.373	0.352	0.392	0.728	0.752

Note: Outcomes are standardized end-of-course grades (math for Columns (1) and (3), and ELA for Columns (2) and (4)) and standardized test scores (Columns (5) and (6)). Each column reports coefficients from a separate regression. All coefficients are multiplied by 10 to ease interpretation. Test windows are defined by the California Education Code and constructed using school days in the spring semester. State tests typically end in weeks 21–22. The absence variables used in this table are counts of absences that occur before, during, and after this test window. All models control for classroom and neighborhood-by-year FE, student characteristics, and total absences in both math and ELA. Standard errors are clustered at the school level and shown in parentheses. + $p < 0.10$ * $p < 0.05$ ** $p < 0.01$.

their analyses of the effects of teacher and student absences, respectively, on student test scores, though they do not benchmark these estimates against effects on course grades.

Columns (3) and (4) of Table 5 re-estimate the course-grade regressions, this time distinguishing between absences that occurred before, during, and after the testing window. For both subjects, absences during and after the testing window (i.e., late in the year) have substantially larger impacts on course grades than earlier absences, though absences in all three time periods have statistically significant impacts. These larger effects likely result from students having less time to make up missed work following absences that occur late in the school year.

Similarly, Columns (5) and (6) of Table 5 conduct the same falsification test for the baseline test-score regressions. These estimates follow a similar pattern as those for course grades before and during the testing window: for both subjects, all absences have statistically significant effects on test scores, though the effects of absences during the testing window are about 5 to 6 times more harmful than absences earlier in the year. Intuitively, the larger effects during the testing window likely reflect missed instructional and preparation time for state tests. However, unlike in the course-grade regressions in Columns (3) and (4), there is no significant effect of post-testing window absences on test scores, which lends further credibility to a causal interpretation of the baseline estimates and is consistent with the results of Gottfried and Kirksey (2017). The effects of absences by timing on both course grades and state tests are plotted side by side in Fig. 3, which makes clear the difference between post-window effects on grades and test scores.

5.3. Probing the identifying assumptions

In this section we probe the identifying assumptions for the baseline "proxy" identification strategy discussed in Section 3.1. The first assumption was that students' preferences for one subject over the other do not endogenously vary over time (and similarly that student-year shocks affect absences in all subjects the same way). While this assumption cannot be directly tested, we now conduct two auxiliary analyses to show that, at least to a first approximation, this is a reasonable assumption that does not drive our main results.

First, this approach exploits within-student-year variation in subject-specific absences. While students' math and ELA absences in a given year are highly correlated (0.85 correlation coefficient), two-thirds of student-year observations have different numbers of math and ELA absences. Why these discrepancies exist is key to the validity of our identification strategies, which assume that such differences are conditionally random (e.g., due to the timing of a dentist appointment or leaving school early due to illness) and not the result of a year-specific preference for one subject over the other that might also create corresponding differences in academic performance.[21] Specifically, the concern is that time-varying subject-specific preferences are present in the idiosyncratic error terms in Eq. (4), perhaps due to subject-specific λ in Eq. (2), which might bias the estimates. We provide suggestive evidence that these differences are conditionally random by showing that absences in *both* subjects respond almost identically to lagged test scores and lagged absences in subject *j*.

The left panel of Appendix Fig. A1 plots both math and ELA absences as a function of lagged math achievement, partialling out lagged ELA achievement. The two plots are nearly identical, suggesting that math and ELA absences differ for reasons unrelated to math ability or preferences (which are loosely captured by lagged math performance). If differences between math and ELA absences were driven by subject-specific preferences or ability, we would expect the ELA plot to be a relatively flat, horizontal line. The fact that both math and ELA absences have nearly the same negative, approximately linear relationship with lagged math scores suggests that absences are generally sticky within students and are associated with negative shocks and poor academic performance in a general as opposed to a subject-specific sense. As shown in the right panel of Appendix Fig. A1, the same is true for the relationship between current math and ELA absences and lagged ELA scores, although the plots are much more flat compared with those using lagged math scores.

Appendix Fig. A2 shows similar patterns in the relation of current absences to lagged absences. Specifically, current math and ELA absences both respond in nearly identical ways to lagged math absences, and similarly for lagged ELA absences. Together, the patterns depicted in Figs. A1, A2 suggest that between-subject absence differences in a given year are not driven by students' interest, ability, or past performance in a specific subject. This is consistent with the similarity between the lag-score and student-FE specifications reported in Table 3 and suggest that changes over time in subject-specific performance or ability do not drive differences in subject-specific absences.

Second, we augment the baseline "proxy model" to directly control for student-year specific factors that might create a student-year preference for one subject over the other. Specifically, we augment the model to account for student–teacher match quality, which is another potentially endogenous source of year-specific, subject-specific preferences that would operate through classroom preferences. We add separate indicators for whether the student and teacher are of the same sex or same race, given that demographically-matched secondary school teachers increase achievement (Dee, 2007; Lim and Meer, 2017) and likely influence attendance as well (Holt and Gershenson, 2019; Tran and Gershenson, 2021; Liu and Loeb, 2021). Appendix Table A1 reports these estimates, where we see that the main results are quite robust to the inclusion of the race- and gender-match indicators. This provides additional evidence that the baseline estimates are not biased by unobserved, year-specific, subject-specific preferences and thus that they represent causal effects.

Finally, Appendix Table A2 shows that the baseline estimates are robust to using absences in other (non-ELA, non-math) classes to proxy for the unobserved student-year shock, including physical education. This is an important result because it suggests that the proxy strategy's redundancy assumption holds, at least to a first approximation. The intuitive reason is that this assumption rules out spillover effects of absences in subject *j* on achievement in subject *k*. At first glance, this assumption seems questionable, as more ELA absences could create more make-up work that takes away time from students' math preparation, or even directly hinder reading comprehension that is useful in math class. However, such spillover effects would likely vary by class, and be vanishingly small in the case of physical education classes. That the five sets of estimates using five different "off-subject" absences, including physical education absences, are all in the neighborhood of the baseline estimate of -0.04 suggests that spillover effects (i.e., failure of the redundancy assumption) are not a practically important concern.[22]

In sum, the checks presented in this section provide suggestive evidence that the identifying assumptions are generally plausible, and while they might not hold exactly, any resulting bias is likely

[21] Time-invariant student preferences for a subject are accounted for by student FE and the results are robust to doing so, as reported in Section 4.1.

[22] The intuition behind this test is similar to an over-identification test that compares the 2SLS estimates generated by different instruments (Hausman, 1978).

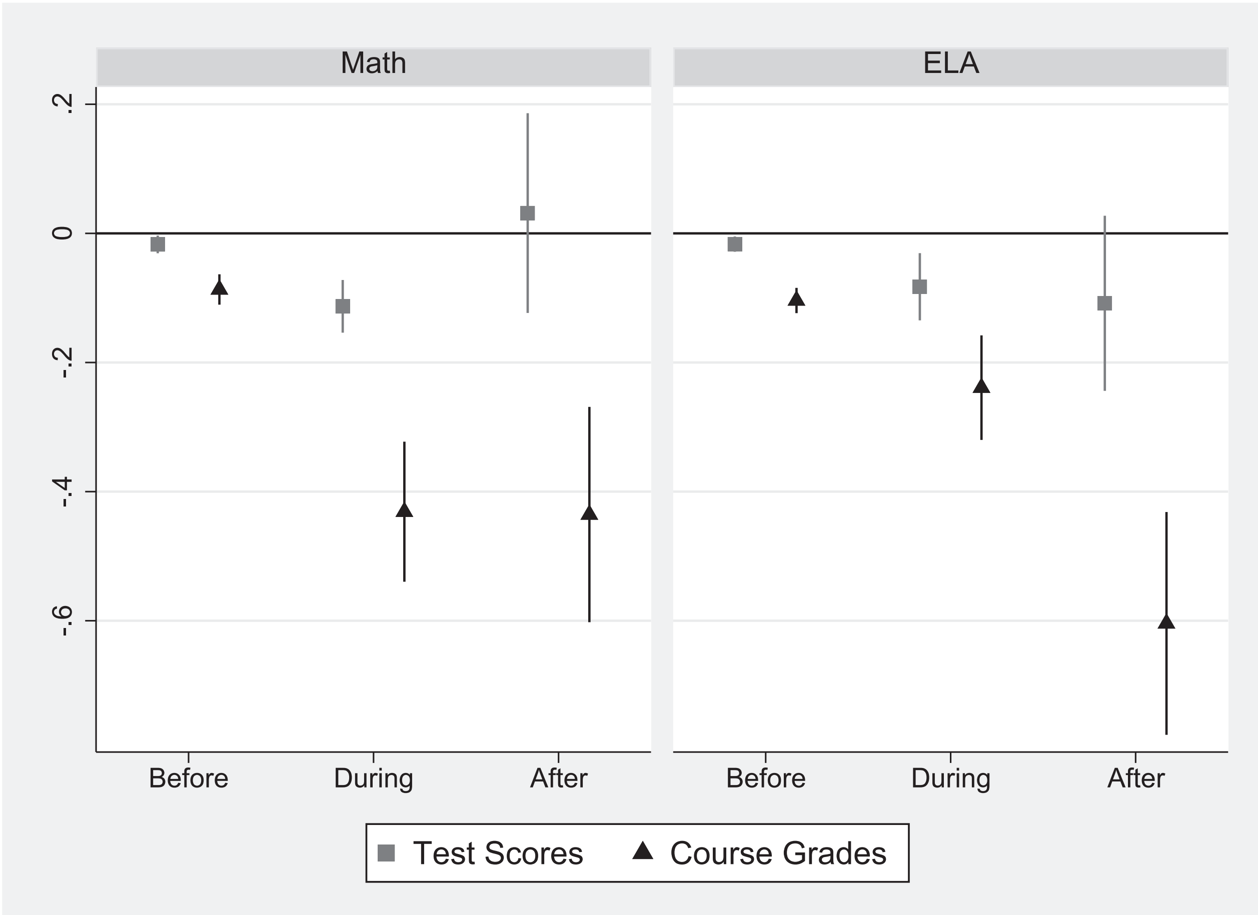

Fig. 3. The Effects of Absences Before, During, and After Test Window on Test Scores and Course Grades. Note: Test windows are defined by the California Education Code and constructed using school days in the spring semester. State tests typically end in weeks 21–22. All models control for total math and ELA absences pre-, during-, and post test window. All coefficients are scaled by a factor of 10 to ease interpretation. Bars show confidence intervals at the 95% level.

small enough to be practically unimportant. Together with the falsification test results presented in Section 5.2, these findings support a causal interpretation of the baseline estimates reported in Section 4.1.

6. Long-run effects

The results presented thus far provide compelling evidence that middle and high school absences harm student learning, as measured by standardized end-of-year tests and course grades. However, test scores and grades are primarily interesting to the extent that they proxy for longer-term outcomes of policy interest, such as high school completion and college going. Short-run effects of schooling inputs on test scores do not always perfectly predict long-run effects, and it is not obvious that all forms of educational attainment would be equally affected by high school absences (e.g., high school graduation, two-year college enrollment, four-year college enrollment). The correlation between high school absences and negative long-run outcomes, such as dropout, drug use, and criminal activity, is well documented (Rumberger and Rotermund, 2012; Hawkins et al., 1998; Henry and Huizinga, 2007; Loeber and Farrington, 2000; Rocque et al., 2017). However, this research is largely correlational. These concerns motivate a direct analysis of the causal relationship between student absences and educational attainment.

Part of the reason that there is scant credible evidence on the long-run impacts of absences is the lack of an obvious research design: we cannot use the proxy strategy described in Section 3.1 because the proxy (total annual absences) is essentially the treatment of interest in this case, nor can we rely on student FE because there is no within-student variation in the outcomes of interest. Accordingly, we adopt a straightforward selection-on-observables strategy in which we regress various measures of educational attainment on either ninth- or tenth-grade total absences, school-by-year FE, neighborhood-by-year FE, and student covariates (including ELL status, demographic indicators, and lagged test scores and grades). The five outcomes are high school graduation, immediate (after graduation) college enrollment, any college enrollment, any enrollment in a four-year college, and any enrollment in a two-year college. These estimates are reported in Table 6.

Panel A of Table 6 shows that ten absences in ninth grade reduce all five measures of educational attainment by about one percentage point, or 2%, and that these effects are statistically significant. Panel B shows similar effects of tenth grade absences, though a subtle difference is that the effect on four-year enrollment is slightly larger than on two-year enrollment, which is intuitive. These modest, but nontrivial, estimates suggest that while much of the raw correlation between absences and long-run outcomes documented elsewhere is driven by selection into absences, there is likely a direct effect of absences as well.

Still, while the school and neighborhood FE and student covariates arguably control for a good amount of selection into absences, there is still some potential selection on unobservables. Accordingly, we also estimate bounds for the causal effects that are based on an assumption about the amount of selection on unobservables into student absences relative to that on observables (Altonji et al.,

Table 6
Impact of absences on long-term outcomes.

	High School Graduation (1)	Immediate College Enrollment (2)	Ever Enrolled in College (3)	Ever Enrolled in 4-Year (4)	Ever Enrolled in 2-Year (5)
A. 9th Grade					
Total Absences	−0.014**	−0.012**	−0.013**	−0.009**	−0.009**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Oster Bound	−0.011	−0.008	−0.009	−0.005	−0.008
Delta	5.392	3.210	3.977	2.236	12.991
Outcome Averages	0.703	0.557	0.619	0.448	0.419
Observations	25,189	25,038	25,038	25,038	25,038
B. 10th Grade					
Total Absences	−0.013**	−0.013**	−0.013**	−0.011**	−0.008**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Oster Bound	−0.011	−0.008	−0.010	−0.005	−0.008
Delta for $\beta = 0$	5.064	2.603	3.766	1.892	69.570
Outcome Averages	0.808	0.634	0.710	0.508	0.488
Observations	31,163	30,941	30,941	30,941	30,941

Note: Each column under each panel shows estimates from a separate regression estimating long-term impact of absences. Each model predicts a long-term outcome of interest using total number of student absences across all subjects (i.e., math, ELA, science, social studies, foreign languages, PE, and other courses) accrued in a given year (Grade 9 in Panel A, Grade 10 in Panel B). All models control for student-level covariates and classroom and neighborhood-year FE. Coefficients, standard errors, and Oster bounds are scaled by a factor of 10 to ease interpretation. Oster bounds are computed based on the assumption that the maximum R^2 is 1.3 times as big as the R^2 from the full model. Sample sizes vary between the first column and Columns (2) through (5), as postsecondary data linkages are conditional on high school graduation and exclude a small number of students who do not graduate from high school. Standard errors are clustered at the school level and shown in parentheses. + p < 0.10 * p < 0.05 ** p < 0.01.

2005; Oster, 2019). Even when assuming a strong degree of selection on unobservables (i.e., that there are equal amounts of selection on both observed and unobserved student characteristics), we still find nonzero, statistically significant negative effects of absences on the long-run outcomes in question. Put differently, there would have to be an implausibly large amount of selection on unobservables (i.e., at least twice as much) as there is on observables to fully explain away the estimated impacts of absences on long-run outcomes; these are the Delta values reported in Table 6. Given that the observables include lagged achievement and school and neighborhood FE, coupled with the modest changes in test score effects seen in Table 3 when the proxy was added to the baseline model, this is extremely unlikely. Thus, the long-run estimates in Table 6 provide arguably compelling evidence that high school absences affect high school completion and college enrollment rates. These effects are fairly constant across types of college enrollment and reinforce a causal interpretation of the main finding that absences harm student achievement and, ultimately, their educational attainment.

7. Conclusion

Using detailed administrative data that track student attendance in each class on each day, this study provides the cleanest causal estimates to date of the immediate and longer-run impacts of secondary school student absences. Our novel approach exploits two distinct yet closely related identification strategies that leverage between-subject variation in student absences to account for unobserved student-year specific shocks. Both approaches yield similar results. Specifically, missing ten class meetings in middle and high school reduces end-of-year test scores by 3–4% of a SD in both math and ELA classes. We then use new developments in the selection-on-observables literature to show that student absences in the ninth and tenth grades affect educational attainment as well, at least among students on the margin of obtaining credentials: ten absences reduce the probability of high school graduation and college enrollment by 1.3 percentage points (or 2%). To our knowledge, this is the first credible evidence of the long-run harms attributable to secondary school student absences and reaffirms the importance of addressing student absenteeism.

Leveraging data on the timing of state standardized tests, we also provide an additional check on the validity of the main results by showing that post-exam absences do not affect exam scores. We then combine this information with data on course grades to probe possible mechanisms driving the results. Because absences later in the year are more detrimental to course grades and exam scores, and in the case of course grades even after the testing window has closed, we can rule out a pure teaching-to-the-test explanation. Instead, it appears that late-year absences are simply harder to make up, either because there is less time to do so or because students are busier and encountering more difficult material later in the year.

Consistent with evidence from elementary schools, we find approximately linear effects of student absences on test scores that are relatively constant across student socio-demographic groups. This linearity suggests that chronic absenteeism indicators, which are widely used in education policy-making, are arbitrary and that using chronic absenteeism rates as an accountability metric misses a large portion of absences that cause substantial learning loss.

These effects are also economically consequential: extrapolating from the teacher-effect estimates in Chetty et al. (2014), eliminating ten math-class absences in a single grade could increase a student's lifetime earnings by about $12,000.[23] Doing so across multiple subjects and multiple years of middle and high school could amount to earnings gains of more than $100,000. Another way to assess the economic significance of these estimates is to use them to infer the learning loss caused by school closures during the COVID-19 pandemic (Kuhfeld et al., 2020).[24] Specifically, the average student missed about 60 days of in-person schooling in the spring of 2020.[25] Given the abrupt nature of school closures and the confusion about re-opening and online-learning plans, it is reasonable to

[23] Chetty et al. (2014) estimate that the SD of teacher effects in middle school is 0.098 for ELA and 0.134 for math. They also estimate that students would gain approximately $39,000 on average in cumulative lifetime income from a one SD improvement in teacher value-added in a single grade. We use teacher value-added in math to provide a rough lower bound for the impact of absences on lifetime income.

[24] Similarly, taken at face value, our estimates suggest that cutting the summer vacation in half (i.e., adding 35 days to the school year) would boost achievement by about 0.14 SD. This is a large effect, given that the average learning gains over the course of ninth grade are about 0.25 SD in math (Bloom et al., 2008). However, this estimate should be interpreted as an upper bound for two general reasons. First, absences are likely costly over and above the lost instructional time due to the coordination problem associated with making up that missed time (Lazear, 2001). Second, the summer vacation counterfactual is not zero, as many students enjoy a variety of experiences that develop their social and human capital (Gershenson, 2013).

[25] The majority of school districts in the U.S. shut down around March 15, 2020. A typical school year ends in mid June, so students missed about three months of instruction or 60 instructional days.

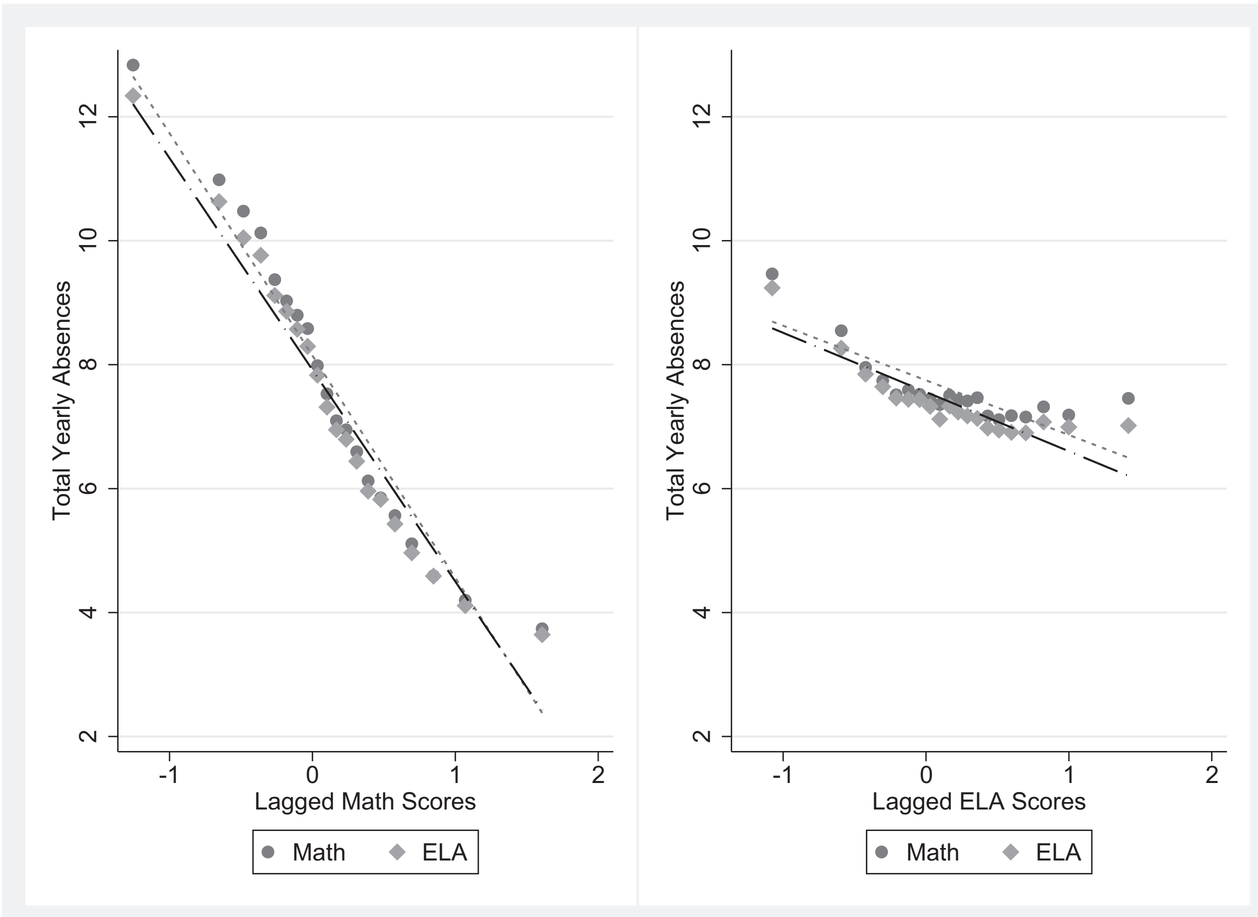

Fig. A1. Binned Scatter Plot of Absences vs. Lagged Test Scores. Note: Data are from the school years 2002–2003 to 2012–2013. Observations are counts of student absences in math and ELA classes in a school year corresponding to binned prior year math or ELA test scores, after partialling out the other subject's test scores.

assume that little to no learning took place on about half, or 30, of those missed days. Extrapolating from our baseline estimates and assuming linearity, this implies a learning loss of about $30 \times 0.004 = 0.12$ test-score SD in both math and ELA and a potential 2 to 3 percentage point (3 to 4%) decrease in the likelihood of high-school graduation and college enrollment.[26]

Credible evidence that absences harm achievement also has implications for racial achievement gaps, which have shrank over the past 40 years but remain sizable (Reardon, 2011), given that there are pronounced racial differences in absence rates (Whitney and Liu, 2017). For example, in our analytic sample Black and Hispanic students score 1.12 and 0.94 SD lower, respectively, on the state math test than their white and Asian peers. Meanwhile, Table 2 shows the average white or Asian student misses seven math classes per year while Black and Hispanic students miss 23 and 17, respectively. While we find no heterogeneity by race in the impact of absences, these stark differences in absence rates likely contribute to achievement gaps. For example, reducing Black and Hispanic student absences to the level of their white and Asian peers would increase the average achievement of Black students by about $16 \times 0.004 = 0.064$ test-score SD and that of Hispanic students by about 0.04 SD. Relative to the raw achievement gaps, absences might explain as much as 6 and 4% of the Black and Hispanic achievement gaps, respectively. While not a silver bullet, these are nontrivial reductions and every opportunity to reduce racial disparities in educational outcomes is worthy of serious consideration.

An exciting development in the field of absence reduction is that numerous interventions have recently been piloted, rigorously evaluated, and brought to scale. Many of these are behaviorally informed and rely on nudges that provide information or reminders of the importance of regular attendance. These interventions are relatively low cost and appear to be very cost effective when compared to the likely lifetime earnings gains associated with improved attendance. For example, one high-frequency text message intervention targeted to parents increased class attendance by 12% at a modest cost of $63 per student (Bergman and Chan, 2021). A similar intervention that aims to correct parents' biased beliefs about their children's absence histories might reduce absences at a cost of $6 per absence at scale, across all grade levels (Rogers and Feller, 2018). Other than behavioral interventions, structural school policies can affect student attendance as well, such as bussing options (Gottfried, 2017) and class size (Tran and Gershenson, 2021). Another way to do this is to formally recognize that teachers affect student attendance and actively develop and reward this dimension of teacher quality (Gershenson, 2016; Liu and Loeb, 2021). Relatedly, efforts to increase diversity and racial representation in the teaching force can close racial disparities in absence rates, as same-race teachers improve Black and Hispanic students' attendance (Holt and Gershenson, 2019; Tran and Gershenson, 2021). Future work should systematically evaluate and compare the costs and benefits of these different approaches for absence reduction, as well as the optimal targeting of these various approaches.

[26] These estimates are roughly consistent with more rigorous projections and survey data on COVID's educational impacts (Kuhfeld et al., 2020; Aucejo et al., 2020; Azevedo et al., 2020).

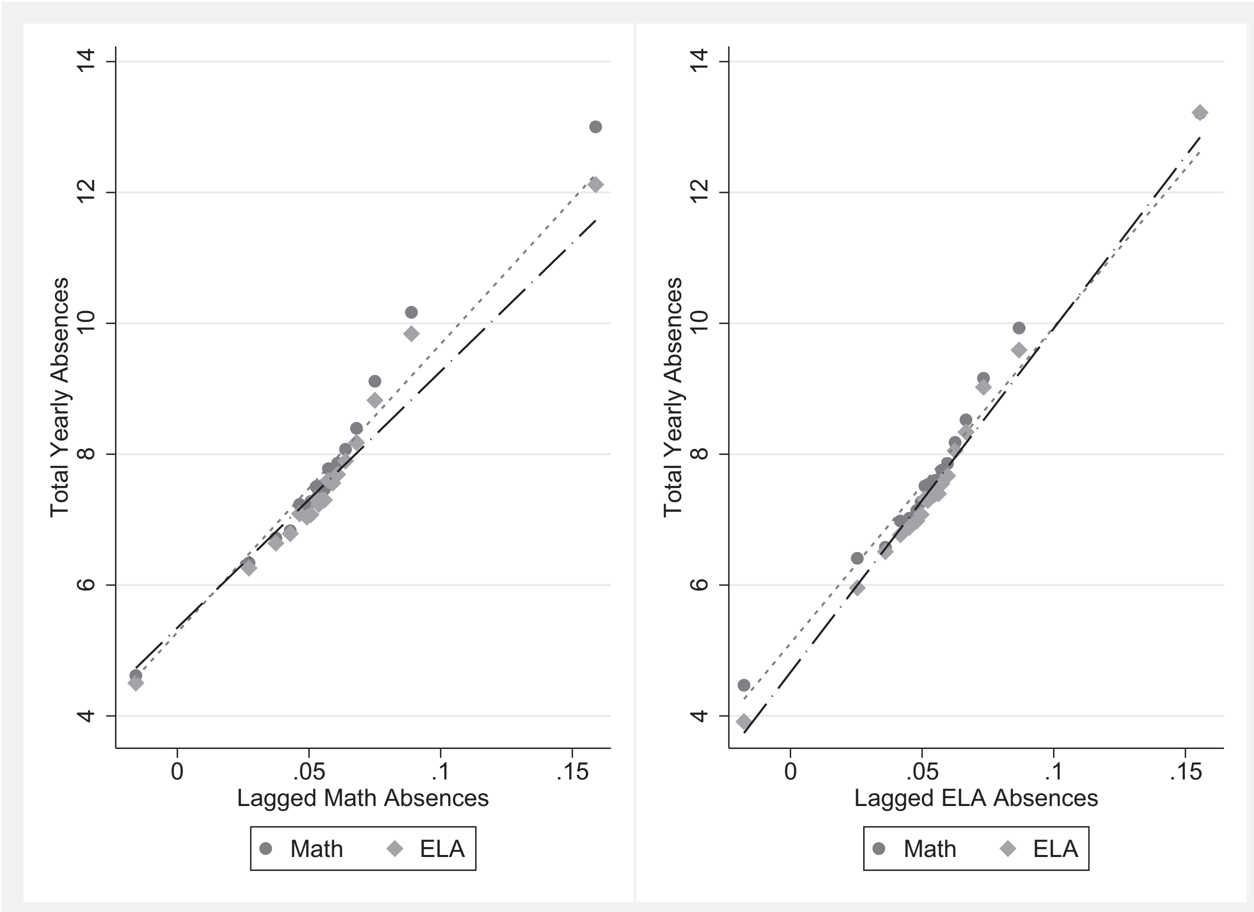

Fig. A2. Binned Scatter Plot of Absences vs. Lagged Absences. Note: Data are from the school years 2002–2003 to 2012–2013. Observations are counts of student absences in math and ELA classes in a school year corresponding to binned prior year math or ELA absences, after partialling out the other subject's annual absences.

Appendix A

.

Table A1
Replication of main results controlling for race/gender match.

	Subject Specific Model		Stacked Model
	Math (1)	ELA (2)	Math + ELA (3)
Subject Absences	−0.042**	−0.040**	−0.028**
	(0.009)	(0.007)	(0.004)
ELA + Math Absences	−0.025**	−0.015**	
	(0.005)	(0.004)	
Race Match	0.163**	−0.019	0.534**
	(0.060)	(0.080)	(0.039)
Gender Match	0.139**	0.025	0.298**
	(0.045)	(0.032)	(0.027)
Classroom FE	X	X	X
Lagged Math & ELA Scores	X	X	
Student Controls	X	X	
Neighborhood-year FE	X	X	
Student-Year FE			X
R^2	0.754	0.721	0.841
Observations	117,450	112,707	333,624

Notes: Each column shows regression outputs from separate models. Results shown are replications of Table 3 controlling for teacher-student race/gender match effects, where race/gender match equals one if the teacher and student are in the same race category (White, Hispanic, Black, and Asian) or gender (Male, Female). Columns (1) and (2) show coefficients on subject-specific absences (Math and ELA, respectively) while Column (3) shows results using data that stack Math and ELA observations. All coefficients are multiplied by 10 to ease interpretation. Student-level controls include both linear and quadratic lagged math and ELA test scores, lagged total absence rate, lagged total suspension days, race, gender, ELL status, disability status, and Special Education status. Classroom- and school-level controls use the same set of control variables as the student level. All models cluster standard errors at the school level. Standard errors are clustered at the school level and shown in parentheses. + p < 0.10 * p < 0.05 ** p < 0.01.

Table A2
Robustness check using off-subject absences.

	Science (1)	Social Studies (2)	Foreign Langs (3)	PE (4)	Other (5)
A. Math					
Math Absences	−0.013+	−0.035**	−0.054**	−0.057**	−0.042**
	(0.008)	(0.009)	(0.008)	(0.007)	(0.006)
Math + X Absences	−0.037**	−0.028**	−0.017**	−0.020**	−0.025**
	(0.004)	(0.004)	(0.003)	(0.004)	(0.003)
R^2	0.721	0.721	0.721	0.721	0.721
Observations	112,711	112,711	112,711	112,711	112,711
B. ELA					
ELA Absences	−0.000	−0.025**	−0.034**	−0.047**	−0.027**
	(0.007)	(0.007)	(0.006)	(0.005)	(0.006)
ELA + X Absences	−0.034**	−0.023**	−0.018**	−0.013**	−0.023**
	(0.005)	(0.003)	(0.003)	(0.003)	(0.003)
R^2	0.755	0.755	0.754	0.754	0.755
Observations	117,453	117,453	117,453	117,453	117,453

Note: Each column and panel show coefficients from a separate regression. The table replicates the results from Table 3 by using one alternative subject at a time, listed in Columns 1 through 5. Panel A shows the impact of absences from math and the total absences from an alternative subject and math on math test scores, while Panel B shows the impact of absences from ELA and the total absences from an alternative subject and ELA on ELA test scores. Standard errors are clustered at the school level and shown in parentheses. + $p < 0.10$ * $p < 0.05$ ** $p < 0.01$.

References

Altonji, Joseph G., Elder, Todd E., Taber, Christopher R., 2005. Selection on observed and unobserved variables: Assessing the effectiveness of Catholic schools. Journal of Political Economy 113 (1), 151–184.

Ashenfelter, Orley, Krueger, Alan, 1994. Estimates of the Economic Return to Schooling from a New Sample of Twins. Am. Econ. Rev. 84 (5), 1157–1173.

Aucejo, Esteban M., Romano, Teresa Foy, 2016. Assessing the effect of school days and absences on test score performance. Economics of Education Review 55, 70–87.

Aucejo, Esteban M., French, Jacob, Ugalde Araya, Maria Paola, Zafar, Basit, 2020. The impact of COVID-19 on student experiences and expectations: Evidence from a survey. Journal of Public Economics 191, 104271.

Autor, David, Figlio, David, Karbownik, Krzysztof, Roth, Jeffrey, Wasserman, Melanie, 2016. School quality and the gender gap in educational achievement. American Economic Review 106 (5), 289–295.

Azevedo, João Pedro, Hasan, Amer, Goldemberg, Diana, Iqbal, Syedah Aroob, Geven, Koen, 2020. Simulating the potential impacts of COVID-19 school closures on schooling and learning outcomes: A set of global estimates. The World Bank.

Bauer, Lauren, Patrick Liu, Diane Whitmore Schanzenbach, and Jay Shambaugh, "Reducing Chronic Absenteeism under the Every Student Succeeds Act," Brookings Institution, 2018.

Bergman, Peter, Chan, Eric W., 2021. Leveraging Parents through Low-Cost Technology The Impact of High-Frequency Information on Student Achievement. Journal of Human Resources 56 (1), 125–158.

Bloom, Howard S., Hill, Carolyn J., Black, Alison Rebeck, Lipsey, Mark W., 2008. Performance trajectories and performance gaps as achievement effect-size benchmarks for educational interventions. Journal of Research on Educational Effectiveness 1 (4), 289–328.

Bound, John, Solon, Gary, 1999. Double trouble: on the value of twins-based estimation of the return to schooling. Economics of Education Review 18 (2), 169–182.

Carrell, Scott E., Maghakian, Teny, West, James E., 2011. A's from Zzzz's? The causal effect of school start time on the academic achievement of adolescents. American Economic Journal: Economic Policy 3 (3), 62–81.

Chetty, Raj, Friedman, John N., Rockoff, Jonah E., 2014. Measuring the impacts of teachers I: Evaluating bias in teacher value-added estimates. American Economic Review 104 (9), 2593–2632.

Chetty, Raj, Friedman, John N., Rockoff, Jonah E., 2014. Measuring the impacts of teachers II: Teacher value-added and student outcomes in adulthood. American Economic Review 104 (9), 2633–2679.

Dee, Thomas S., 2007. Teachers and the gender gaps in student achievement. Journal of Human Resources 42 (3), 528–554.

Dundar, Afet, Shapiro, Doug, 2016. The National Student Clearinghouse as an integral part of the national postsecondary data infrastructure. National Student Clearinghouse Research Center 20, 1–18.

Easton, John Q., Johnson, Esperanza, Sartain, Lauren, 2017. The predictive power of ninth-grade GPA. University of Chicago Consortium on School Research, Chicago, IL.

Gershenson, Seth, 2013. Do summer time-use gaps vary by socioeconomic status? American Educational Research Journal 50 (6), 1219–1248.

Gershenson, Seth, 2016. Linking teacher quality, student attendance, and student achievement. Education Finance and Policy 11 (2), 125–149.

Gershenson, Seth, Jacknowitz, Alison, Brannegan, Andrew, 2017. Are student absences worth the worry in US primary schools?. Education Finance and Policy 12 (2), 137–165.

Gershenson, Seth, Holt, Stephen B., Papageorge, Nicholas W., 2016. Who believes in me? The effect of student–teacher demographic match on teacher expectations. Economics of Education Review 52, 209–224.

Goodman, Joshua, "Flaking out: Student absences and snow days as disruptions of instructional time," Technical Report, National Bureau of Economic Research 2014.

Gottfried, Michael A., 2009. Excused versus unexcused: How student absences in elementary school affect academic achievement. Educational Evaluation and Policy Analysis 31 (4), 392–415.

Gottfried, Michael A., 2010. Evaluating the relationship between student attendance and achievement in urban elementary and middle schools: An instrumental variables approach. American Educational Research Journal 47 (2), 434–465.

Gottfried, Michael A., 2011. The detrimental effects of missing school: Evidence from urban siblings. American Journal of Education 117 (2), 147–182.

Gottfried, Michael A., 2017. Linking getting to school with going to school. Educational Evaluation and Policy Analysis 39 (4), 571–592.

Gottfried, Michael A. and E. Hutt, Addressing Absenteeism, Harvard Education Press, 2019.

Gottfried, Michael A., Kirksey, J., 2017. When students miss school: The role of timing of absenteeism on students test performance. Educational Researcher 46 (3), 119–130.

Hausman, Jerry A., 1978. Specification tests in econometrics. Econometrica: Journal of the Econometric Society, 1251–1271.

Hawkins, J David, Todd Herrenkohl, David P. Farrington, Devon Brewer, Richard F. Catalano, and Tracy W. Harachi, "A review of predictors of youth violence." 1998.

Heissel, Jennifer A., Norris, Samuel, 2018. Rise and shine the effect of school start times on academic performance from childhood through puberty. Journal of Human Resources 53 (4), 957–992.

Henry, Kimberly L., Huizinga, David H., 2007. Truancys effect on the onset of drug use among urban adolescents placed at risk. J. Adolesc. Health 40 (4), 358–3e9.

Herrmann, Mariesa A., Rockoff, Jonah E., 2012. Worker absence and productivity: Evidence from teaching. Journal of Labor Economics 30 (4), 749–782.

Holt, Stephen B., Gershenson, Seth, 2019. The impact of demographic representation on absences and suspensions. Policy Studies Journal 47 (4), 1069–1099.

Jackson, C. Kirabo, 2014. Teacher quality at the high school level: The importance of accounting for tracks. Journal of Labor Economics 32 (4), 645–684.

Jacob, Brian A., Lovett, Kelly, 2017. Chronic absenteeism: An old problem in search of new answers. Brookings Institution, Washington, DC.

Kirksey, J. Jacob, 2019. Academic Harms of Missing High School and the Accuracy of Current Policy Thresholds: Analysis of Preregistered Administrative Data From a California School District. AERA Open 5 (3). 2332858419867692.

Kuhfeld, Megan, Soland, James, Tarasawa, Beth, Johnson, Angela, Ruzek, Erik, Liu, Jing, 2020. Projecting the potential impact of COVID-19 school closures on academic achievement. Educational Researcher 49 (8), 549–565.

Lazear, Edward P., 2001. Educational production. Q. J. Econ. 116 (3), 777–803.

Lim, Jaegeum, Meer, Jonathan, 2017. The impact of teacher–student gender matches random assignment evidence from South Korea. Journal of Human Resources 52 (4), 979–997.

Liu, Jing, Loeb, Susanna, 2021. Engaging Teachers: Measuring the Impact of Teachers on Student Attendance in Secondary School. The Journal of Human Resources 56 (2), 343–379.

Loeber, Rolf, Farrington, David P., 2000. Young children who commit crime: Epidemiology, developmental origins, risk factors, early interventions, and policy implications. Dev. Psychopathol. 12 (4), 737–762.

Oster, Emily, 2019. Unobservable selection and coefficient stability: Theory and evidence. Journal of Business & Economic Statistics 37 (2), 187–204.
Pope, Nolan G., 2016. How the time of day affects productivity: Evidence from school schedules. Rev. Econ. Stat. 98 (1), 1–11.
Reardon, Sean F., 2011. The widening academic achievement gap between the rich and the poor: New evidence and possible explanations. Whither Opportunity? 1, 91–116.
Rice, Jennifer King, Croninger, Robert G., Roellke, Christopher F., 2002. The effect of block scheduling high school mathematics courses on student achievement and teachers' use of time: implications for educational productivity. Economics of Education Review 21 (6), 599–607.
Rocque, Michael, Jennings, Wesley G., Piquero, Alex R., Ozkan, Turgut, Farrington, David P., 2017. The importance of school attendance: Findings from the Cambridge study in delinquent development on the life-course effects of truancy. Crime & Delinquency 63 (5), 592–612.
Rogers, Todd, Feller, Avi, 2018. Reducing student absences at scale by targeting parents misbeliefs. Nature Human Behaviour 2 (5), 335–342.
Rumberger, Russell W. and Susan Rotermund, "The relationship between engagement and high school dropout," in "Handbook of research on student engagement," Springer, 2012, pp. 491–513.
Tran, Long and Seth Gershenson, "Experimental Estimates of the Student-Attendance Production Function," Educational Evaluation & Policy Analysis, 2021.
Whitney, Camille R., Liu, Jing, 2017. What we're missing: A descriptive analysis of part-day absenteeism in secondary school. AERA Open 3 (2). 2332858417703660.
Wooldridge, Jeffrey M., 2010. Econometric analysis of cross section and panel data. MIT press.